



Deep Learning Models for Aspect-Based Sentiment Analysis Task: A Survey Paper

Sarsabene Hammi^(✉), Souha Mezghani Hammami^(✉),
and Lamia Hadrich Belguith^(✉)

ANLP Research Group, MIRACL Lab, FSEGS, University of Sfax, Sfax, Tunisia
hsarsabene@gmail.com, souha.hamhami@ihecs.usf.tn,
lamia.belguith@fsegs.usf.tn

Abstract. Due to the significant increase in the volume of data shared on the web, Aspect-Based Sentiment Analysis (ABSA) has become essential. This task ensures a detailed sentiment analysis. It identifies firstly the aspect terms (e.g., price, food, etc.) and then classifies their sentiment polarity as positive, negative, or neutral. Many approaches have been used to treat this task including the machine learning-based approach, the rule-based approach, etc. However, with the important increase in the content of the internet, these approaches became relatively unable to analyze this volume of information, resulting in the emergence of the deep learning-based approach which is the subfield of the machine learning-based approach.

Recently many researchers used the deep learning-based approach to address the ABSA. This paper provides a summary of the deep learning models that have been developed for ABSA, as well as a survey of studies that have employed these models to address different subtasks of the ABSA task. Finally, we discuss the implications of our work and potential avenues for future research.

Keywords: Aspect-based Sentiment Analysis · Aspect Extraction · Aspect Sentiment Classification · Deep Learning

1 Introduction

The Sentiment Analysis (SA) task is a subfield of Natural Language Processing (NLP) that involves using artificial intelligence and information retrieval techniques to identify and extract opinions, emotions, and other subjective information from text. The principal objective of SA is to gain insights into the general sentiment of a group of people toward a specific topic. This research direction has gained significant attention in academia and industry due to its ability to assist in marketing decision-making and track shifts in customer opinions on various subjects, including the medical domain (such as the COVID pandemic). Previous studies divided the SA task into three main levels. The first one is the document-level sentiment analysis, which is focused on identifying the general opinion of a document (such as a tweet, review, or article), and determining whether it is positive, negative, or neutral. The sentence-level sentiment analysis is the second type. It

concentrates on identifying the opinion of individual sentences within documents. The third one is Aspect-Based Sentiment Analysis (ABSA), which offers a more detailed and precise analysis. This analysis involves two tasks. The first one is the Aspect Extraction (AE) task, which identifies the aspect terms of a certain entity. The second one is the Aspect Sentiment Classification (ASC) task and it tends to determine the opinion related to aspects identified in the AE task. Taking as an example the following comment, the ABSA determines first the aspects “*camera*” and “*fingerprint reader*”. Then classifies sentiment polarity related to these aspects. In this case, positive sentiment is given to the aspects “*camera*” and “*fingerprint reader*” of the entity “*smartphone*”.

The camera is very good. The fingerprint reader works well.

Early studies categorized the approaches of ABSA task into four main approaches. The rule-based approach, machine learning-based approach, deep learning-based, and hybrid approach. In this survey paper, we focused only on the research papers that employed deep learning models to tackle the ABSA task.

The organization of the rest of this paper is as follows: In Sect. 2, we provide a broad summary of the deep learning models that are utilized in ABSA. Section 3, and 4 summarizes the different studies proposed respectively for AE and ASC tasks. Section 5 describes research papers that treat simultaneously the AE and ASC. Section 6, discusses different models used and gives statistics about the most performant model in the ABSA task. Section 7 concludes this study.

2 Deep Learning Models

Deep learning is a category of machine learning that draws inspiration from the organization and operation of the mind, particularly its neural networks. Deep learning models use multiple layers of artificial neural networks to learn and make decisions. Each layer receives information from the preceding layer and exploits it to produce another information beneficial to the classification.

The deep learning models have many advantages in comparison to the other machine learning models. Firstly, these models are able to improve their performance over time through the process of training. During this process, the network’s biases and weights are adjusted based on the accuracy of predictions. One of the significant advantages of these models is their ability to enhance their performance gradually by means of training. During this process, the network’s weights and biases are adjusted based on the precision of its forecasts, resulting in improved accuracy over time. Also, the deep learning models are self-adaptive. They are able to adapt to the data and find features on their own, without the need for the functional or distributional form of the model to be defined beforehand. So, the ability to learn, adapt to the data, and improve the performance without explicit programming makes deep learning models particularly well-suited for sentiment analysis tasks, where the features and relationships in the data may be complex and difficult to specify in advance.

Several deep-learning models have been employed in the ABSA. The upcoming subsections will offer a detailed account of each deep learning model.

2.1 Classical Recurrent Neural Network Model

The Recurrent Neural Network (RNN) model is a popular deep learning model. Recently, it has been greatly utilized in several NLP tasks, including the ABSA task [1]. This big use can be explained by the good results achieved by this model in the treatment of sequential data. The principal idea of the RNN model lies in the treatment of tokens that compose the input in a sequential manner. The classical RNN model follows the mechanism of forward propagation. At instant t , the RNN model feeds tokens of input sequence (X) into a neural network architecture which is composed of nodes interconnected between them. Then, assuming the connection between the different tokens of the sequence, the RNN model uses the outputs of previous nodes (h_{t-1}) to estimate the value of current node (h_t). The final node contains all information about the tokens appearing with the aspect, treated from left to right. This information is included, finally in an output layer to predict the label. The RNN model's architecture is presented in Fig. 1.

In addition, RNN model can follow also the mechanism of backward propagation. A backward RNN works in the same manner as a forward RNN but in the opposite direction: from right to left (the prediction is made from the end towards the beginning of the sequence). It aims in recovering the information from the next node to calculate the value of the current node at instant t . An RNN model that used the mechanism of forward propagation, in addition to backward propagation is named the bidirectional RNN model (Bi-RNN). This model is very performant and surpasses the classical RNN model thanks to its performant architecture.

The RNN model is a performant model that uses the contextual information between words, nevertheless, it still suffers from problems such as the vanishing gradients problem and the long-term dependencies learning.

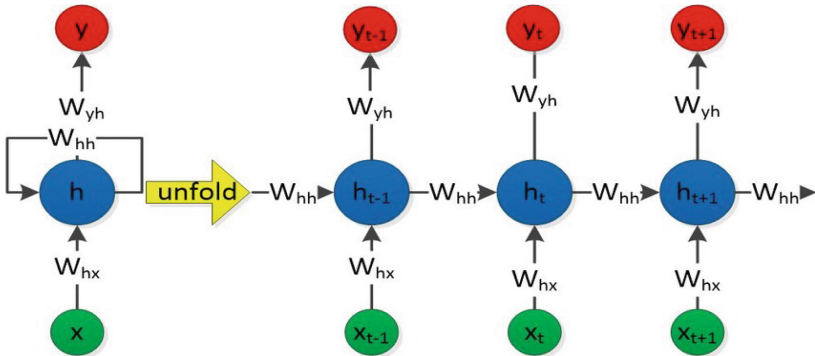


Fig. 1. RNN model's architecture.

2.2 Long Short-Term Memory Model

To address some of the problems faced by the RNN model, a variant called the Long Short-Term Memory (LSTM) was introduced [2]. As we mentioned above these problems are related mainly to the vanishing gradients and long-term dependencies learning.

To address these problems, researchers replaced the classic recurrent hidden unit with a memory unit. The LSTM unit is composed of a central node, containing the internal state (or memory) of the unit, and three gates. The input gate decides whether the cell's state must be updated or not. The output gate detects the next hidden state's value. The forget gate chooses which information should be ignored. Like RNNs, LSTMs follow the mechanism of forward propagation and treat the input sequence in a unidirectional manner. The LSTM model's architecture is presented in Fig. 2.

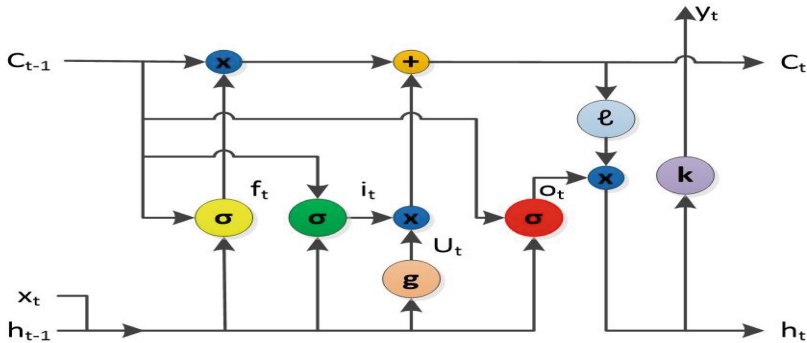


Fig. 2. LSTM unit's architecture.

Besides the LSTM model, we can find also a bidirectional LSTM (Bi-LSTM) model. It determines a word's label by leveraging the information coming from the previous units (forward propagation) and next units (backward propagation).

Though the LSTM model solved in part the issues related to the RNN model, it still has drawbacks, including the lengthy training period and the complex recurrent unit architecture.

2.3 Gated Recurrent Units Model

The Gated Recurrent Units (GRU) model was also proposed to solve issues related to the RNN model. This model came as an improvement to the RNN model and it deals with the issues of vanishing gradients and the learning of extended dependencies. The GRU comprises a cell state and two gates [3]. The update is responsible for deciding whether or not to update the hidden unit's state. The reset gate determines the degree to which prior information should be discarded. The GRU model follows the mechanism of forward propagation. Also, it can be improved by a bidirectional GRU (Bi-GRU) model that treats the input sequence according to forward propagation and backward propagation. Figure 3 presents the GRU model's architecture.

This model is a very performant model which is characterized by a simple architecture compared to the LSTM model, however, it has some drawbacks such as the small learning ability.

2.4 Convolution Neural Networks Model

The Convolution Neural Networks (CNN) model specializes in multi-layered networks. It is probably used when the input is structured according to a grid (e.g. an image). These networks were inspired by the work of [4] on the visual cortex of animals.

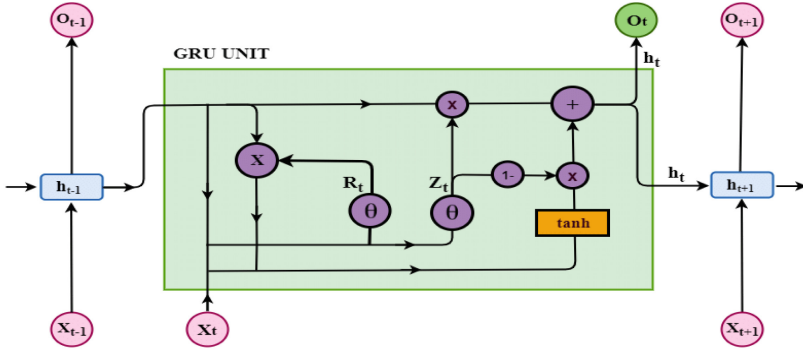


Fig. 3. The GRU unit’s architecture.

The CNN model is initially introduced by [5] to treat the task of forms recognition and other tasks such as image classification, and character recognition. Recently, this model proved its relevance in NLP tasks and especially for tasks related to text classification. The CNN model’s architecture is presented in Fig. 4. It contains mainly four types of layers. The convolution layer is a principal layer in the CNN model and it multiplies the matrix representation M by another called convolution matrix (or filter) to produce a feature map. The pooling layer reduces the feature map’s dimensions while keeping only relevant information. The fully-connected layer refers to a neural network wherein neurons are related. The output layer predicts the adequate class.

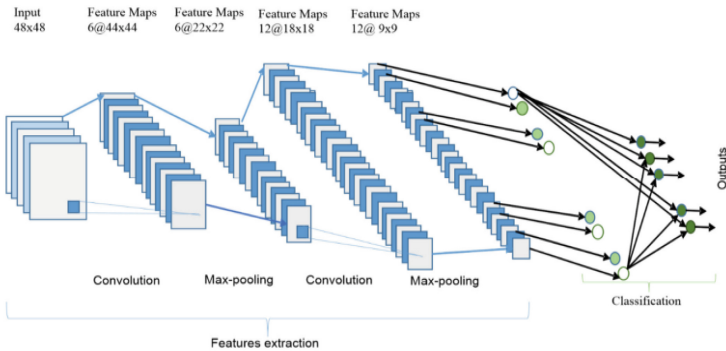


Fig. 4. The CNN model’s architecture.

2.5 Bidirectional Encoder Representations from Transformers Model

The Bidirectional Encoder Representations from Transforms (BERT) model is a cutting-edge NLP model presented by Google [6]. This model is based on transformers and uses attention mechanisms to acquire contextual relationships among words within a text. A vast corpus of information is utilized to train BERT. Upon being trained on an extensive data corpus, BERT is fine-tuned to accomplish diverse NLP tasks such as sentiment analysis and named entity recognition.

One of the key advantages of BERT is that it is bidirectional, meaning that it takes into account contextual information of the target word when making predictions, as opposed to traditional models that only consider the context to the left. This allows BERT to perform better on many NLP tasks, particularly those that require an understanding of the context in which a word appears.

BERT has achieved impressive outcomes on an extensive array of NLP benchmarks and has been widely adopted in the NLP community. It has also inspired the development of several related models, including RoBERTa, which builds upon the original BERT architecture and exhibits enhanced performance on some tasks. The BERT model's architecture is presented in Fig. 5.

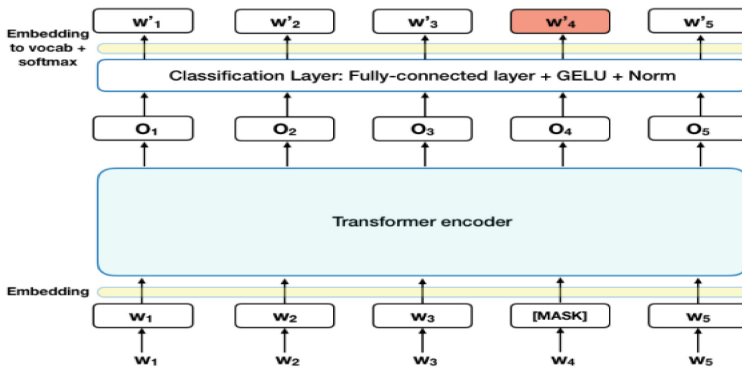


Fig. 5. The BERT model's architecture.

3 Deep Learning Models for AE Task

As we discussed above, the Aspect Extraction (AE) task is an important task in any ABSA-related work. It goals to identify the aspects within a text. Recently many studies have focused on the AE task only (without treating the ASC task). In this section, we are interested in the studies that have used deep learning models in the AE task. Among these studies, we mention the work of [7] that used an RNN model for identifying aspects existing in SemEval-2014 dataset. To achieve this work, the authors transformed, first, every word within the dataset to a word vector. These vectors were built based on Amazon Embeddings and SENNA Embeddings systems. Afterward, these vectors were used

to create new context vectors that take into consideration the contextual dependencies between words. Finally, the RNN model used the constructed vectors (word embeddings vectors and context vectors) and the linguistic features' vectors to determine aspects. In the study presented in [8], an RNCRF model that combines the RNN model and the machine learning model: conditional random fields (CRF) was proposed. To implement the RNCRF model, authors constructed a dependency-tree RNN (DT-RNN) architecture. This architecture produces, for each word in the dataset, a high-level representation that takes into consideration the dependency relations between words in the dataset. These representations are included in a subsequent step in the CRF model to predict aspects.

[9] proposed a model based on LSTM for the extraction of aspects related to question answering (ASC-QA) task. In this study, authors constructed first a human-annotated benchmark dataset. After that, they proposed a Reinforced Aspect-relevant Word Selector (RAWS) model in order to select the aspect-relevant words. These selected words were incorporated in a subsequent step in the Reinforced Bidirectional Attention Network (RBAN) architecture to extract the aspect terms. This architecture treats the semantic matching problem in the QA text pair and enhances the learning algorithm. It incorporates both a bidirectional attention mechanism and a reinforcement learning (RL) component. By using a bidirectional attention mechanism, the model can identify both the aspect and its corresponding context. The RL component assists in enhancing the model's ability to comprehend the connections existing amidst the aspect and its context. This work was implemented using the LSTM model.

Other studies think that the Bi-LSTM model is more efficient than the LSTM model. For that reason, many studies used this model in their AE-related methods. Among them, we mention the work of [10] which used a multi-layers Bi-LSTM model in their work. [10] assumed that the incorporation of information about both words and clauses inside the Bi-LSTM model can enhance mainly the process of aspects detection. To realize this work, [10] segmented first the sentences into clauses. After that, they incorporated contextual vectors into the word-level aspect-specific attention layer. This layer exploits the contextual information and outputs new vectors that contain the degree of importance of each word in a given clause. These newly produced vectors were fed in the clause-level attention layer to detect which clauses are important in the dataset. Finally, the Bi-LSTM model used the vectors extracted by the word-level aspect-specific attention and clause-level attention layers to forecast the aspect terms in the dataset. This method achieved good results (68.50% F-measure). In another work, [11] enhanced the AE task using the Bi-LSTM model and Bidirectional Dependency Tree Conditional Random Field Framework (BiDTreeCRF). So, the authors constructed first a bidirectional dependency tree network (BiDTree) in order to detect the dependency relationships among the words. Afterward, they included the output of BiDTree into a Bi-LSTM model to detect the global syntactic context of each word. Finally, the outcomes produced by the preceding steps were fed into a CRF model to predict aspect terms.

[12] considered the AE as a sequence-to-sequence (seq2seq) task and proposed a Bi-GRU-based model. This model uses a seq2seq learning-based architecture that takes into consideration the meaning of sentences and labels in the extraction of aspect terms. At this level, the model takes as input word embeddings and predicts as output the label of a given word.

Other research papers preferred to combine multiple deep learning algorithms in order to perform the aspect extraction task. Among them, we mention the work of [13] that combined the Bi-GRU, CNN, and BERT models to extract the aspect terms. In this work, [13] proposed a new framework named pre-trained language embedding-based contextual summary and multi-scale transmission network (PECSMT). This framework consists of three units. The pre-training language model embedding unit generates contextualized embeddings using a BERT model. The multi-scale transmission network unit uses the multi-scale CNNs and Bi-GRU models to extract the sequential features. The contextual summary unit creates a contextualized representation of words. This model with its three units achieved good results and succeeded in the extraction of aspect terms. [14] introduced a new information-augmented neural network (IAAN) model. This model integrated informative information about the words surrounding the aspect term in order to extract the dynamic word sense. This model involves several layers. The initial layer is a contextualized embedding layer and it uses the BERT model to create contextualized word embeddings. The second layer is an encoder named MCRN and it uses the GRU model to capture the sequential data and bidirectional distant dependencies. The third layer is a decoder and it uses the GRU model to decode the encoding representations in order to predict the aspect terms.

[15] suggested a synchronous double-channel recurrent network (SDRN) model to achieve the Aspect Opinion Pair Extraction (AOPE) task. To realize this model, [15] employed first the word embeddings of BERT to learn the words' contextual semantics. Subsequently, they used this contextual information and the CRF model to detect the

Table 1. Summary of Aspect Extraction Studies.

Study	Domain	Dataset Size	Model	Available	F-measure
[7]	Restaurant Laptop	3841 3845	RNN	yes yes	81.56% 77.27%
[8]	Restaurant Laptop	3841 3845	RNN CRF	yes yes	84.93% 78.42%
[9]	Products	12076	LSTM	no	64.80%
[10]	Restaurant	3841	Bi-LSTM	yes	68.50%
[11]	Restaurant Laptop	8613 3845	Bi-LSTM CRF	yes yes	84.83% 80.57%
[12]	Restaurant Laptop	2000 3045	Bi-GRU	yes yes	75.14% 80.31%
[13]	Laptop Restaurant	3845 3841	BERT Bi-GRU CNN	yes yes	85.08% 88.13%
[14]	Laptop Restaurant	3845 3841	BERT GRU	yes yes	87.49% 88.40%
[15]	Laptop Restaurant	3845 3841	BERT CRF	yes yes	83.67% 89.49%

aspect and opinion terms. Table 1 contains an overview of the different AE-related studies presented.

4 Deep Learning Models for ASC Task

This part of the study gives a summary of the studies that have treated only the ASC task using deep learning models. Among them, we mention the work of [16] that proposed an RNN-based method to perform the ASC for Arabic hotels' reviews dataset. To achieve this method, [16] incorporated into the RNN model lexical, word, syntactic, morphological, and semantic features. These features enhanced significantly the effectiveness of the suggested model. [17] proposed a Target-Connection LSTM (TC-LSTM) model to detect the sentiment polarity towards the aspect terms. This model leverages the semantic relatedness between the aspect term and its context. The TC-LSTM model takes as input words embedding and aspect vectors. These aspect vectors contain information about the contextual words related to a given aspect term. This model achieved competitive results and overrode the other benchmarks, even though syntactic parsers and external sentiment lexicons were used. Similarly, to [17, 18] exploited the context of aspects and suggested an Attention-based LSTM with Aspect Embedding (ATAE-LSTM) to perform the ASC task. This model utilizes attention weights, computed using word embeddings, to capture the information associated with the aspect term. [19] proposed a hierarchical LSTM model. This model takes advantage of the interdependencies between sentences in a review to achieve the ASC task. The achieved outcomes proved the efficacy of the suggested model. Although this model didn't use hand-crafted features, it surpassed the other state-of-art models and achieved competitive results.

[20] enhanced the ASC task using linguistic regularizers and the CNN model. In this work, [20] incorporated in the CNN model two regularizers which are the Coordinating Conjunctions Regularizer (CCR) and Adversative Conjunctions Regularizer (ACR). These regularizers ameliorated the introduced model and achieved good results for the SemEval-2014 dataset. Table 2 summarizes the different studies presented for the ASC task.

Table 2. Summary of Aspect Sentiment Classification Studies.

Study	Domain	Dataset Size	Model	Available	F-measure
[16]	Hotel	340000	RNN	no	-
[17]	Twitter	-	LSTM	yes	65.9%
[18]	Restaurant Laptop	3841 3845	LSTM	yes yes	-
[19]	Multi-Domain	-	LSTM	yes	80.23%
[20]	Restaurant	3841	CNN	yes	-

5 Deep Learning Models for AE and ASC Tasks

In this section, we are interested in the studies that simultaneously treat the ABSA's two tasks: the AE and the ASC tasks. Amongst them, we mention the study of [21] which assumed that the treatment of the AE task and the ASC simultaneously is more beneficial than the treatment of each one of them separately. For that, [21] suggested a DOER (Dual crOss-sharEd RNN) framework to extract the aspects as well as the sentiment polarity. DOER consists mainly of two units: the dual RNN unit and the crOss-sharEd unit. These two units work together and enhance the AE and ASC tasks by using embeddings related to domain-specific and general-purpose.

Other studies used the LSTM model in the AE and ASC task. Amongst them, we mention the study of [22]. [22] used in their work two LSTM models. These models detect the latent relations between opinion words and aspects using a DMI (multi-hop dual memory interaction) mechanism. This mechanism is very performant and succeeded in the realization of both tasks. [23] ameliorated the work of [22] and proposed two-stacked Bi-LSTMs units. The first unit detects the unified tags (the aspect terms and their sentiment). The second unit enhances the prediction performance of the first unit. [24] proposed also a Bi-LSTM model for the identification of aspects in addition to their corresponding sentiment. This model leverages the dependency among aspects and sentiment words using a Bi-LSTM model and Biaffine score. [25] presented a Bi-LSTM-CRF model. This model incorporates first the contextualized representations of words in the Bi-LSTM in order to identify the aspects. These representations are used to detect the interactions between words. In a subsequent step, the binary classifier CRF is employed to assign a sentiment for each aspect. The model's performance was assessed on the products dataset and it demonstrated good performance in comparison to the baseline models. [26] leveraged semantic and syntactic relationships between the opinion words, and aspect terms, and suggested a model based on LSTM to tackle the AE and ASC tasks. For that, [26] modeled first the syntactic and semantic dependencies between words using Graph Neural Networks (GNN). Then, they incorporated these dependencies features and word embeddings into the LSTM model with the aim of identifying the aspects and the corresponding sentiment associated with them.

[27] adopted a CNN-based architecture to perform the ABSA task including the AE task and ASC task. In this work, [27] took advantage of the relatedness between the ABSA-related tasks and proposed an IMN network (interactive multi-task learning network). This network ensures information passing between the ABSA tasks (AE, ASC, etc.) using a common group of latent variables. This proposed method proved its efficiency in the AE and ASC tasks and achieved good results.

[28] proposed a BERT-based model to accomplish the ASC and AE tasks. This proposed method used a framework for joint learning of multiple tasks, where only one model was trained to perform these two related tasks simultaneously. The model can share information between the two tasks, allowing it to enhance the two tasks' performance. In addition, they used two independent BERT layers to extract features belonging to global and local contexts. Such features enhanced significantly the AE and ASC tasks and led to favorable outcomes for both Chinese and English reviews. In another work [29] introduced a deep contextualized relation-aware network (DCRAN) model. This BERT-based model was designed to be context-aware, taking into account

the words and phrases that appear both before and after the aspect or sentiment being identified. This model was also designed to be relation-aware, considering explicitly and implicitly the contextual information between aspects and their sentiments. This proposed method enhanced the current ABSA-related works. Also, [30] focused on the dependencies between opinion and aspect terms and suggested a BERT-based model in order to achieve the AE and ASC tasks.

[31] presented a GRU-based model and Memory Network to address the aspect extraction and aspect-based sentiment analysis tasks. Firstly, [31] pretreated the dataset. Subsequently, they augmented the GRU model with the inclusion of word vectors to extract aspect terms. Finally, they used a Memory Network to classify the sentiment of aspects. The Memory Network takes into consideration the dependence between the aspect terms.

Other studies combined several deep learning models to achieve the AE and ASC tasks. Among these studies, we mention the study of [32] which used the Bi-LSTM-CNN-based model for extracting aspects and classifying their sentiment. In this study, [32] employed a model that realizes multiple related tasks simultaneously. This model surpassed the other benchmarks related to the ABSA in English and Hindi languages. [33] introduced a method for unified aspect-based sentiment analysis. This method used a collaborative learning approach, in which the CNN model is trained to perform the AE and ASC tasks. During the training phase, this model uses the word embeddings provided by the BERT model to create useful vectors. In addition, this method takes into consideration the relationships between aspects, as sentiment towards one aspect can often be influenced by sentiment towards another. [34] presented a method based on the CNN-BERT model. This method employed an interactive architecture, in which a syntactic parser was employed to identify the syntactic dependencies between words in the text. This information was used to guide the aspect identification and classification process. The method also utilized dependency syntactic knowledge, which refers to the relationships between words in a syntactic parse tree, leading to enhanced accuracy for both identifying aspects and classifying sentiments. [35] suggested a Bi-LSTM-BERT model to identify the aspect-sentiment triplets. This model contains a neural network architecture called an Explicit Interaction Network (EIN). This architecture was designed to capture the relationships between different words in the text and then used it to identify aspects and the sentiment expressed towards them. The EIN architecture contains multiple layers that work together through explicit attention mechanisms, allowing the model's ability to concentrate on specific input parts and incorporated context from other parts of the input when making predictions.

Table 3 contains an overview of the different studies already presented for the ABSA task.

6 Discussion

This part of the study compares the deep-learning models utilized to solve ABSA-related tasks. We take into consideration the studies described in Sect. 5 (15 studies) that treat both the AE and the ASC tasks. Figure 6 presents the F-measure values achieved by the deep-learning models in the ABSAs' studies. The obtained results proved that the

BERT model achieved the highest F-measure values. This model has achieved a good performance in a wide range of ABSA-related studies for many reasons. Firstly, the BERT model underwent pre-training on a massive corpus of data, which allows it to understand the context and words' meaning in a sentence. This pre-training can be fine-tuned on specific tasks, which improves performance. Secondly, the BERT was trained in order to understand the word's context by looking at the words that come before and after it. Also, it is a fine-tunable model, which means that it can be easily fine-tuned on specific tasks, even with limited annotated data. All these advantages make the BERT model performant and suitable for the ABSA task.

Table 3. Summary of Aspect-Based Sentiment Analysis Studies.

Study	Domain	Dataset Size	Model	Available	F-measure
[21]	Twitter	3223	RNN	yes	51.37%
	Restaurant	3841		yes	72.78%
	Laptop	3845		yes	60.35%
[22]	Multi-Domain	9359	LSTM	yes	-
[23]	Laptop	2931	Bi-LSTM	yes	57.90%
	Restaurant	9738	RNN	yes	58.90%
[24]	Laptop	1487	Bi-LSTM	yes	45.05%
	Restaurant	4550		yes	54.82%
[25]	Laptop	1435	Bi-LSTM	yes	51.33%
	Restaurant	4536	CRF	yes	61.55%
			BERT		
[26]	Laptop	1456	Bi-LSTM	yes	52.01%
	Restaurant	4533		yes	64.06%
[27]	Laptop	1903	CNN	yes	68.26%
	Restaurant	3859		yes	57.40%
[28]	Multi-Domain	41640	BERT	yes	88.14%
[29]	Restaurant	4190	BERT	yes	72.42%
[30]	Laptop	1453	BERT	yes	58.58%
	Restaurant	4536		yes	64.75%
[31]	Product	1180	GRU	no	67.00%
[32]	Product	26044	Bi-LSTM	yes	74.60%
			CNN		
[33]	Laptop	3048	CNN	yes	63.40%
	Restaurant	4359	BERT	yes	70.43%
[34]	Laptop	3048	CNN	yes	63.04%
	Restaurant	3044	BERT	yes	67.48%
[35]	Laptop	1903	Bi-LSTM	yes	61.60%
	Restaurant	5497	CNN	yes	63.77%

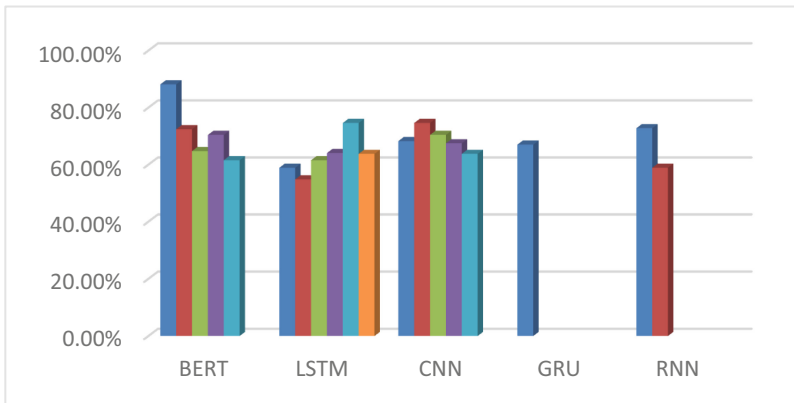


Fig. 6. The distribution of F-measure values achieved through the deep-learning models.

7 Conclusion

This survey paper gives a comprehensive review about different research that utilized deep learning models in solving the ABSA task. We first provided an overview concerning the deep learning models used in the achievement of ABSA tasks, including the RNN (Recurrent Neural Network), LSTM (Long-short term memory), etc. After that, we summarized and explained the studies that treated each one of the ABSA sub-tasks independently of the other subtask: the AE subtask and the ASC subtask. Also, we presented the studies that treat both tasks. Finally, we discussed the models used in the ABSA studies. The obtained results showed that the best performances have been obtained by the BERT.

Our future work intends to provide a survey with a detailed analysis of other ABSA-related studies that have used the linguistic knowledge approach, machine learning-based approach, and hybrid approach. We will also provide an overview of the dataset used in this field. In addition, we will compare different approaches by mentioning the advantages and disadvantages of each approach.

References

1. Rumelhart, D., Hinton, G.E., Williams, R.J.: Learning representations by back-propagating errors. *Nature* **323**, 533–536 (1986)
2. Graves, A.: Long short-term memory. In: *Supervised Sequence Labeling with Recurrent Neural Networks*, pp. 37–45 (2012)
3. Cho, K., Van, B., Bahdanau, D., Bengio, Y.: On the properties of neural machine translation: encoder-decoder approaches. arXiv preprint [arXiv:1409.1259](https://arxiv.org/abs/1409.1259) (2014)
4. Hubel, D.H., Torsten, N.W.: Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.* **160**(1), 106 (1962)
5. Fukushima, K., Sei, M.: Neocognitron: a self-organizing neural network model for a mechanism of visual pattern recognition. In: Amari, S.I., Arbib, M.A. (eds.) *Competition and Cooperation in Neural Nets*, pp. 267–285. Springer, Heidelberg (1982). https://doi.org/10.1007/978-3-642-46466-9_18

6. Devlin, J., Chang, M., Lee, K., Toutanova, K.: BERT: pre-training of deep bidirectional transformers for language understanding. arXiv preprint [arXiv:1810.04805](https://arxiv.org/abs/1810.04805) (2018)
7. Liu, P., Joty, S., Meng, H.: Fine-grained opinion mining with recurrent neural networks and word embeddings. In: Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, pp. 1433–1443 (2016)
8. Wang, W., Pan, S.J., Dahlmeier, D., Xiao, X.: Recursive neural conditional random fields for aspect-based sentiment analysis. arXiv preprint [arXiv:1603.06679](https://arxiv.org/abs/1603.06679) (2016)
9. Wang, J., et al.: Aspect sentiment classification towards question-answering with reinforced bidirectional attention network. In: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pp. 3548–3557 (2019)
10. Wang, J., et al.: Aspect sentiment classification with both word-level and clause-level attention networks. In: IJCAI, pp. 4439–4445 (2018)
11. Luo, H., Li, T., Liu, B., Wang, B., Unger, H.: Improving aspect term extraction with bidirectional dependency tree representation. *IEEE/ACM Trans. Audio Speech Lang. Process.* **27**, 1201–1212 (2019)
12. Ma, D., Li, S., Wu, F., Xie, X., Wang, H.: Exploring sequence-to-sequence learning in aspect term extraction. In: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pp. 3538–3547 (2019)
13. Feng, C., Rao, Y., Nazir, A., Wu, L., He, L.: Pre-trained language embedding-based contextual summary and multi-scale transmission network for aspect extraction. *Procedia Comput. Sci.* **174**, 40–49 (2020)
14. Liu, N., Shen, B.: Aspect term extraction via information-augmented neural network. *Complex Intell. Syst.*, 1–27 (2022)
15. Chen, S., Liu, J., Wang, Y., Zhang, W., Chi, Z.: Synchronous double-channel recurrent network for aspect-opinion pair extraction. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pp. 6515–6524 (2020)
16. Al-Smadi, M., Al-Ayyoub, M., Jararweh, Y., Qawasmeh, O.: Deep recurrent neural network for aspect-based sentiment analysis of Arabic hotels reviews. *J. Comput. Sci.*, 386–393 (2018)
17. Tang, D., Qin, B., Feng, X., Liu, T.: Effective LSTMs for target-dependent sentiment classification. In: COLING, pp. 3298–3307 (2016)
18. Wang, Y., Huang, M., Zhu, X., Zhao L.: Attention-based LSTM for aspect-level sentiment classification. In: EMNLP, pp. 606–615 (2016)
19. Ruder, S., Ghaffari, P., Breslin, J.G.: A hierarchical model of reviews for aspect-based sentiment analysis. In: Conference on Empirical Methods in Natural Language Processing, *ACL*, pp. 999–1005 (2016)
20. Zeng, D., Dai, Y., Li, F., Wang, J., Sangaiah, A.K.: Aspect based sentiment analysis by a linguistically regularized CNN with gated mechanism. *J. Intell. Fuzzy Syst.* **36**(5), 3971–3980 (2019)
21. Luo, H., Li, T., Liu, B., Zhang, J.: DOER: dual cross-shared RNN for aspect term-polarity co-extraction. arXiv preprint [arXiv:1906.01794](https://arxiv.org/abs/1906.01794) (2019)
22. Li, Z., Li, X., Wei, Y., Bing, L., Zhang, Y., Yang, Q.: Transferable end-to-end aspect-based sentiment analysis with selective adversarial learning. arXiv preprint [arXiv:1910.14192](https://arxiv.org/abs/1910.14192) (2019)
23. Li, X., Bing, L., Li, P., Lam, W.: A unified model for opinion target extraction and target sentiment prediction. In: Proceedings of the AAAI Conference on Artificial Intelligence, pp. 6714–6721 (2019)
24. Zhang, C., Li, Q., Song, D., Wang, B.: A multi-task learning framework for opinion triplet extraction. arXiv preprint [arXiv:2010.01512](https://arxiv.org/abs/2010.01512) (2020)
25. Xu, L., Li, H., Lu, W., Bing, L.: Position-aware tagging for aspect sentiment triplet extraction. arXiv preprint [arXiv:2010.02609](https://arxiv.org/abs/2010.02609) (2020)

26. Chen, Z., Huang, H., Liu, B., Shi, X., Jin, H.: Semantic and syntactic enhanced aspect sentiment triplet extraction. arXiv preprint [arXiv:2106.03315](https://arxiv.org/abs/2106.03315) (2021)
27. He, R., Lee, W.S., Ng, H.T., Dahlmeier, D.: An interactive multi-task learning network for end-to-end aspect-based sentiment analysis. arXiv preprint [arXiv:1906.06906](https://arxiv.org/abs/1906.06906) (2019)
28. Yang, H., Zeng, B., Yang, J., Song, Y., Xu, R.: A multi-task learning model for Chinese-oriented aspect polarity classification and aspect term extraction. *Neurocomputing* **419**, 344–356 (2021)
29. Oh, S., et al.: Deep context-and relation-aware learning for aspect-based sentiment analysis. arXiv preprint [arXiv:2106.03806](https://arxiv.org/abs/2106.03806) (2021)
30. Huang, L., et al.: First target and opinion then polarity: enhancing target-opinion correlation for aspect sentiment triplet extraction. arXiv preprint [arXiv:2102.08549](https://arxiv.org/abs/2102.08549) (2021)
31. Ismet, H.T., Mustaqim, T., Purwitasari, D.: Aspect based sentiment analysis of product review using memory network. *Sci. J. Inf.* **9**, 73–83 (2022)
32. Liu, Q., Liu, B., Zhang, Y., Kim, D.S., Gao, Z.: Improving opinion aspect extraction using semantic similarity and aspect associations. In: *Thirtieth AAAI Conference on Artificial Intelligence* (2016)
33. Chen, Z., Qian, T.: Relation-aware collaborative learning for unified aspect-based sentiment analysis. In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (2020)
34. Liang, Y., Meng, F., Zhang, J., Chen, Y., Xu, J., Zhou, J.: A dependency syntactic knowledge augmented interactive architecture for end-to-end aspect-based sentiment analysis. *Neurocomputing* **454**, pp 291–302 (2021)
35. Wang, P., et al.: Explicit interaction network for aspect sentiment triplet extraction. arXiv preprint [arXiv:2106.11148](https://arxiv.org/abs/2106.11148) (2021)